

# Inherent Risk and Uncertainty of Self-Similar Sizing Scales in Agile Software Development

IT-CAST, Wednesday, September 15<sup>th</sup>, 2021, 1400 EDT  
Mr. Peter J. Braxton, Technomics, Inc.

*“Does This T-Shirt Make My Estimate Look Big?”*

# Abstract

*In Agile estimating and planning, self-similar scales such as Planning Poker, Fibonacci Numbers, and T-Shirt Sizing are commonly used to assess the size of software development efforts. We start with the simple question of "What if we are off by one [size in either direction]?" Since such estimates typically rely on Expert Judgment, this supposition is not unreasonable. Exploring various discrete and continuous probability distributions, including the Uniform, Triangular, and Lognormal, this paper derives closed-form solutions for expected growth from the point estimate (i.e., Risk) and the accompanying Uncertainty in the form of a coefficient of variation (CV), and generalizes to any choice of confidence level. It considers the trade-offs in the ratio of these scales, with smaller ratios (like Planning Poker) offering finer gradations but an increased chance of being "wrong," while larger ratios (like T-Shirt Sizing) sacrifice granularity for accuracy. Not only do Fibonacci numbers approach the Golden Ratio (approximately 1.618), but they also have the intuitive advantage that any given size is the sum of the two preceding sizes. Drawing on the author's earlier work ("Understatement of Risk and Uncertainty by Subject Matter Experts," SCEA, 2011), this paper proposes improvements to the traditional elicitation process by replacing abstract numbers on the self-similar scale with analogous data points. Finally, it tests the accuracy of forecasts made using both scales for knowable and future quantities.*

# Meet The Briefer



*Peter Braxton is a Subject Matter Expert at Technomics, performing cost and risk analysis for a number of federal clients. He has played integral roles in the development of both the Software Resources Data Report (SRDR) and the BCF 250 Applied Software Cost Estimating course at Defense Acquisition University (DAU). The inaugural Vice President for Professional Development of the International Cost Estimating and Analysis Association (ICEAA) and multiple Educator of the Year winner, he has shown a long-standing commitment to knowledge sharing within the community. His current research interests include leveraging detailed Agile and DevOps data in forecasting program cost.*

[PBraxton@technomics.net](mailto:PBraxton@technomics.net)

(703) 944-3114 (mobile)

# Acknowledgments

*The author wishes to thank Ken Rhodes, Alex Wekluk, and Dave Brown of Technomics for their review of the presentation and excellent suggestions to improve the clarity of the ideas presented.*

# Outline

- T-Shirt Sizing part 1: *“Double or Half?”*
- T-Shirt Sizing part 2: *“Skew You! From Discrete to Continuous”*
- T-Shirt Sizing part 3: *“‘Alpha’ is for ‘Confidence’”*
- Planning Poker and Fibonacci Numbers: *“The Ratio is Golden”*
- Granularity vs. Accuracy: *“Better to be approximately right than precisely wrong”*
- Testing Efficacy of Self-Similar Scales: *“Empiricists R Us”*
- Improving Expert Judgment: *“Analogizing the Sliding Scale”*
- Road Ahead

# Poll #1

- Which is your primary role in the Agile Software Development process?
  - Estimating
  - Planning
  - Development
  - Testing
  - Management
  - Other

# Sizing Approaches – Definitions

- T-Shirt Sizing
- Planning Poker
- Fibonacci Numbers
- Story Points
- Function Points (FP)
- Simplified Function Points (SiFP)
- Source Lines of Code (SLOC)

# Poll #2

- Which of the following sizing approaches have you used in Agile Software Development?
  - T-Shirt Sizing
  - Planning Poker
  - Fibonacci Numbers
  - Story Points
  - Function Points (FP)
  - Simplified Function Points (SiFP)
  - Other (e.g., SLOC, ESLOC)



# The Basic Idea – Double or Half?!

- In the basic *Who Wants To Be a Millionaire* game, the dollar value (approximately) doubles for each question
  - \$1,000 and \$32,000 are “safe” plateaus
- Beyond \$32,000, the contestant is faced with a choice:
  - Walk with the amount already earned, *or*
  - Go for the next question (“double”) *but*
  - Risk losing all but the \$32K
- For the \$64,000 Question – see what they did there?! – the losing side of the bet is precisely “half”

\$32,000 (10 of 15) - Not Timed		Half
David E. Kelley's production company credit at the end of his TV shows features an old woman exclaiming what?		
• A: Good night	• B: Go away	
• C: Bye-bye	• D: You stinker	

\$64,000 (11 of 15) - Not Timed		Earned
In 2000, what city offered a popular new license plate protesting "Taxation Without Representation"?		
• A: San Juan, PR	• B: Boston, MA	
• C: Washington, DC	• D: Austin, TX	

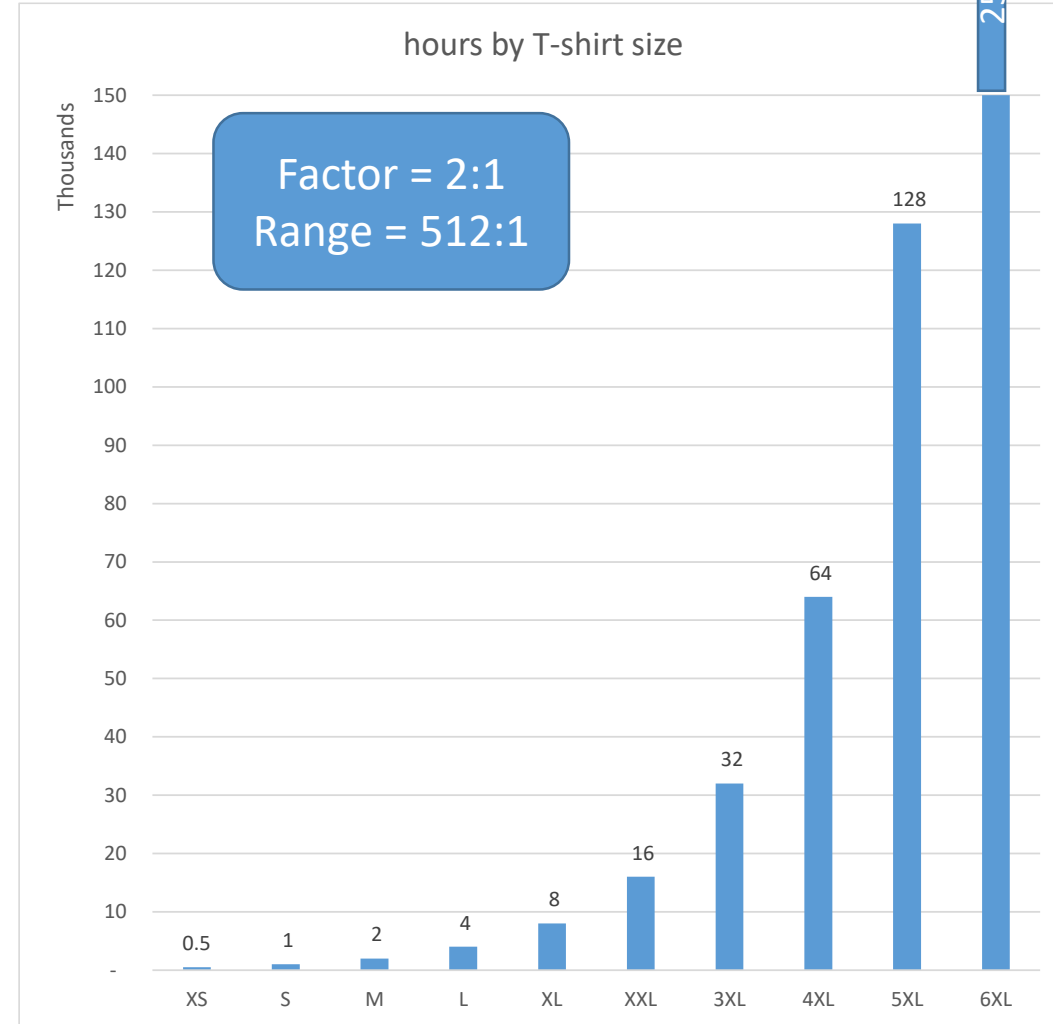
\$125,000 (12 of 15) - Not Timed		Double
What flower comes in a variety of flower types such as spoon, pompon and spider?		
'Phone-a-Friend' and '50:50' lifelines used		
• A: Chrysanthemum	• B: Geranium	
• C: Peony	• D: Rose	

Peter had no clue, so he decided to call his friend Dick, who failed to give an answer within the allotted time. Therefore, he decided to use his 50:50, eliminating B and D. He decided to go with A: Chrysanthemum and won \$125,000.

[https://millionaire.fandom.com/wiki/Peter\\_Braxton](https://millionaire.fandom.com/wiki/Peter_Braxton)

# T-Shirt Sizing Risk – Introduction

- T-Shirt Sizing is purposefully an exponential scale (aka logarithmic)
  - Similar to the use of Fibonacci numbers and “planning poker” in Agile
  - Other common logarithmic scales include Richter (earthquakes) and Decibel (sound)
- Going-in Risk position is that SME assessments could very easily be off by one T-shirt size in either direction
- Straightforward math leads to growth percentages and CVs under various distributional assumptions



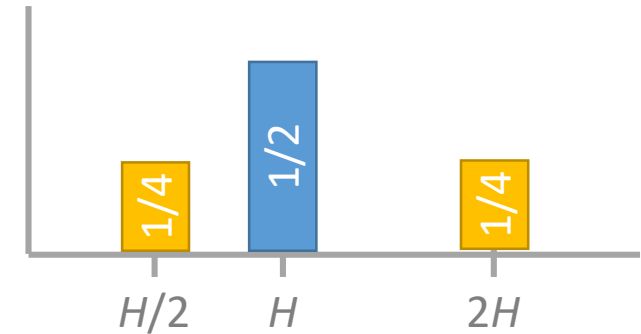
# T-Shirt Sizing Risk – General Framework

- Premise: A variation of the “double-or-half” thought experiment establishes a specific probability distribution
- Risk: Compute the **mean** of the probability distribution
  - Compare to the original point estimate ( $H$  hours) to establish a Cost Growth Factor (CGF), and equivalent **percent growth** (on average)
- Uncertainty: Compute the **variance** of the probability distribution
  - Compare standard deviation to the original point estimate (“pseudo CV”) and estimate with growth to determine Coefficient of Variation (**CV**)
- Refinements:
  1. From discrete to *continuous* outcomes
  2. Incorporating degree of *confidence*
  3. Adjusting *beyond* “double-or-half” based on confidence

# T-Shirt Sizing Risk – Discrete

- Assume a Discrete distribution:

- Most Likely =  $H$  hours, with a probability of  $1/2$
- Max =  $2H$  hours, with a probability of  $1/4$
- Min =  $H/2$  hours, with a probability of  $1/4$



Coin flip #1: right or wrong  
Coin flip #2: high or low

- Mean is expected value: 
$$\sum_i x_i p_i = (1/4)(H/2) + (1/2)(H) + (1/4)(2H) = \frac{9H}{8} = \left(1 + \frac{1}{8}\right)H$$

- CGF = 1.125, or **12.5%** growth over point estimate

- Variance is expected value of square less square of expected value:

$$\sum_i x_i^2 p_i - \left[ \sum_i x_i p_i \right]^2 = (1/4)(H^2/4) + (1/2)(H^2) + (1/4)(4H^2) - \left[ \frac{9H}{8} \right]^2 = \frac{25H^2}{16} - \frac{81H^2}{64} = \left[ \frac{\sqrt{19}}{8} H \right]^2$$

- CV = **48.43%**

# Uniform Distribution

- Assume a Uniform distribution:

- Max =  $2H$  hours (next largest T-shirt size)
- Min =  $H/2$  hours (next smallest T-shirt size)

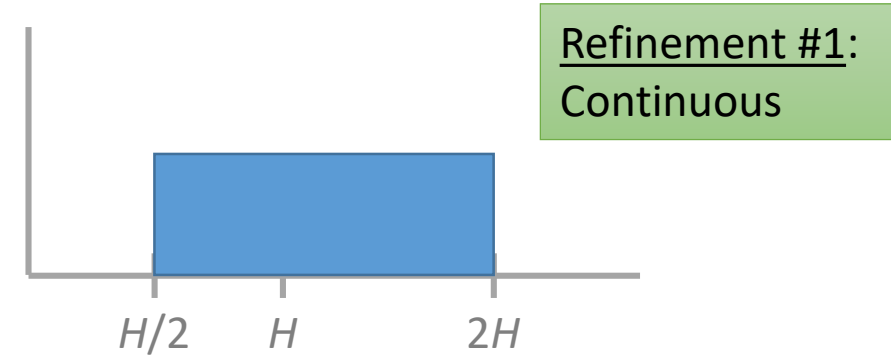
- Mean is average of Min/Max:  $\frac{H/2 + 2H}{2} = \frac{5H}{4} = \left(1 + \frac{1}{4}\right)H$

- CGF = 1.25, or **25.0%** growth over point estimate

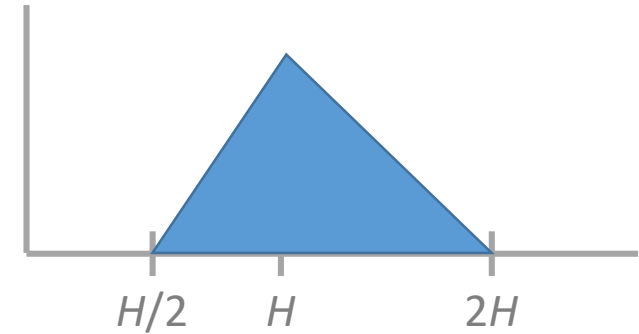
- Variance is range squared / 12:

$$\frac{(2H - H/2)^2}{12} = \frac{9H^2}{4 \cdot 12} = \left[\sqrt{3} \cdot \frac{H}{4}\right]^2 = \left[\frac{\sqrt{3}}{4}H\right]^2$$

- CV = **34.64%**



# Triangular Distribution



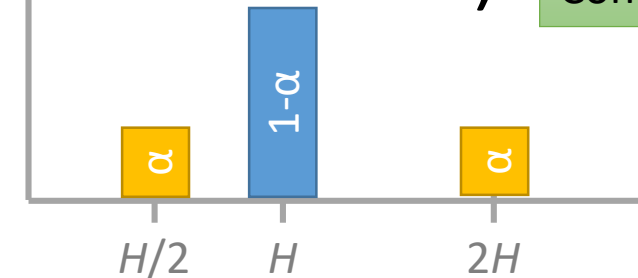
- Assume a Triangular distribution:
  - Most Likely =  $H$  hours (assessed T-shirt size)
  - Max =  $2H$  hours (next largest T-shirt size)
  - Min =  $H/2$  hours (next smallest T-shirt size)
- Mean is average of Min/ML/Max: 
$$\frac{H/2 + H + 2H}{3} = \frac{7H}{6} = \left(1 + \frac{1}{6}\right)H$$
  - CGF = 1.167, or **16.7%** growth over point estimate
- Variance is sum of squares less sum of pairwise products / 18:
 
$$\frac{(H/2)^2 + H^2 + (2H)^2 - H^2/2 - H^2 - 2H^2}{18} = \frac{7H^2/4}{18} = \frac{7H^2}{2 \cdot 36} = \left[\sqrt{\frac{7}{2}} \cdot \frac{H}{6}\right]^2 = \left[\frac{\sqrt{14}}{12}H\right]^2$$
  - CV = **26.73%**

# T-Shirt Sizing Risk – Discrete (Confidence)

Refinement #2:  
Confidence

- Assume a Discrete distribution:

- Most Likely =  $H$  hours, with a probability of  $1-\alpha$
- Max =  $2H$  hours, with a probability of  $\alpha/2$
- Min =  $H/2$  hours, with a probability of  $\alpha/2$



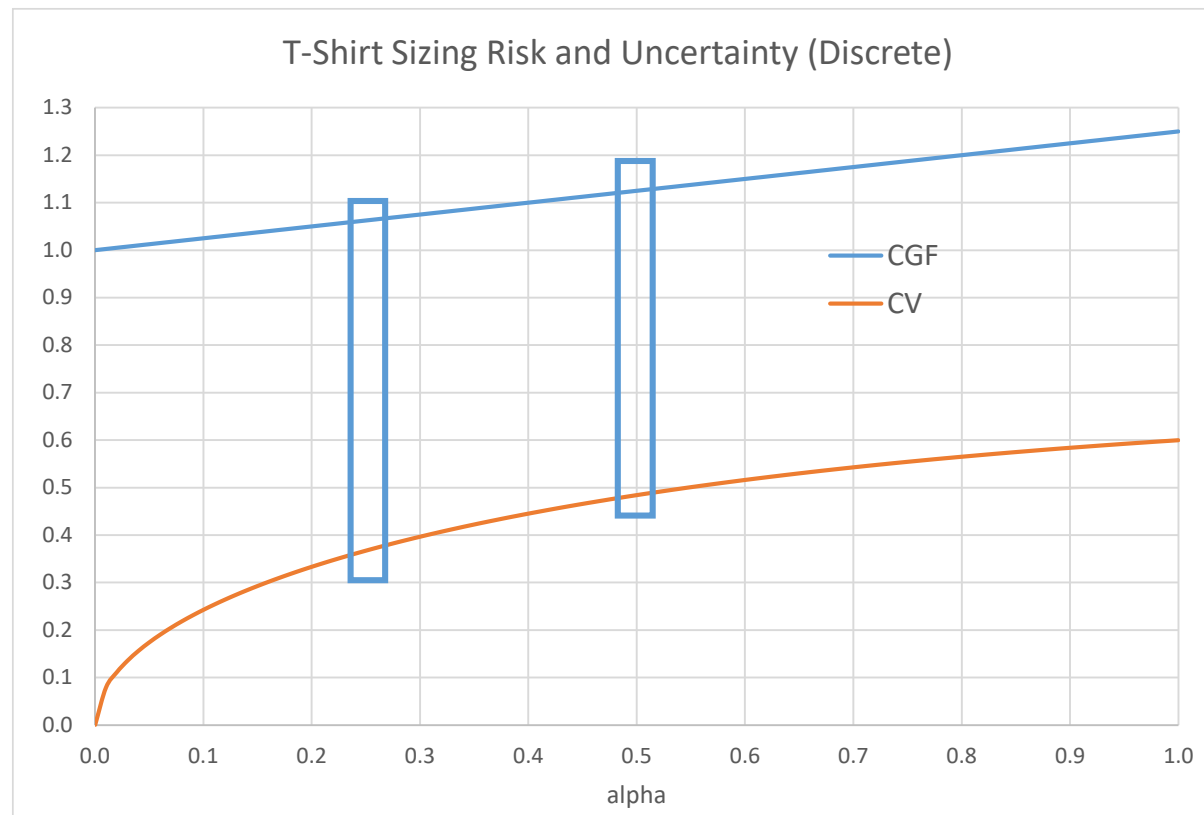
In previous example,  
 $\alpha = 1/2$

- Mean is expected value:  $\sum_i x_i p_i = (\alpha/2)(H/2) + (1-\alpha)(H) + (\alpha/2)(2H) = \left(1 + \frac{\alpha}{4}\right)H$ 
  - CGF =  $1+(\alpha/4)$ , or  $\alpha/4$  growth over point estimate
- Variance is expected value of square less square of expected value:

$$\begin{aligned} \sum_i x_i^2 p_i - \left[ \sum_i x_i p_i \right]^2 &= (\alpha/2) \left( H^2/4 \right) + (1-\alpha)(H^2) + (\alpha/2)(4H^2) - \left[ \left( 1 + \frac{\alpha}{4} \right) H \right]^2 = \\ &= \left( 1 + \frac{9\alpha}{8} \right) H^2 - \left( 1 + \frac{\alpha}{2} + \frac{\alpha^2}{16} \right) H^2 = \frac{10\alpha - \alpha^2}{16} H^2 = \left[ \frac{\sqrt{10\alpha - \alpha^2}}{4} H \right]^2 \quad CV = \frac{\sqrt{10\alpha - \alpha^2}}{4 + \alpha} \end{aligned}$$

# T-Shirt Sizing Risk – Discrete (Illustrated)

- Graph illustrates range between always right ( $\alpha=0$ ) and always wrong ( $\alpha=1$ ), with a coin flip to determine low or high
  - Max growth is 25%
  - Max CV is 60%





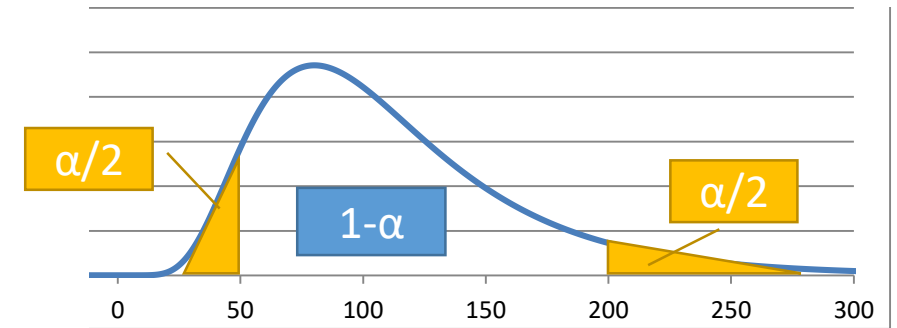
# T-Shirt Sizing Risk – Lognormal

Refinement #3:  
Adjustment

- Assume a Lognormal distribution:
  - Median =  $H$  hours, with a probability of  $1-\alpha$  between  $H/2$  and  $2H$
  - Right tail  $> 2H$  hours, with a probability of  $\alpha/2$
  - Left tail  $< H/2$  hours, with a probability of  $\alpha/2$
- Confidence interval of related normal is:  $(\ln H - \ln 2, \ln H, \ln H + \ln 2)$ 
  - So that  $\Phi^{-1}(1 - \alpha/2) = \frac{\ln 2}{\sigma}$
- Mean of the lognormal is:  $e^{\mu + \frac{\sigma^2}{2}}$ 
  - With a CGF of  $e^{\frac{\sigma^2}{2}} = \sqrt{1 + CV^2}$

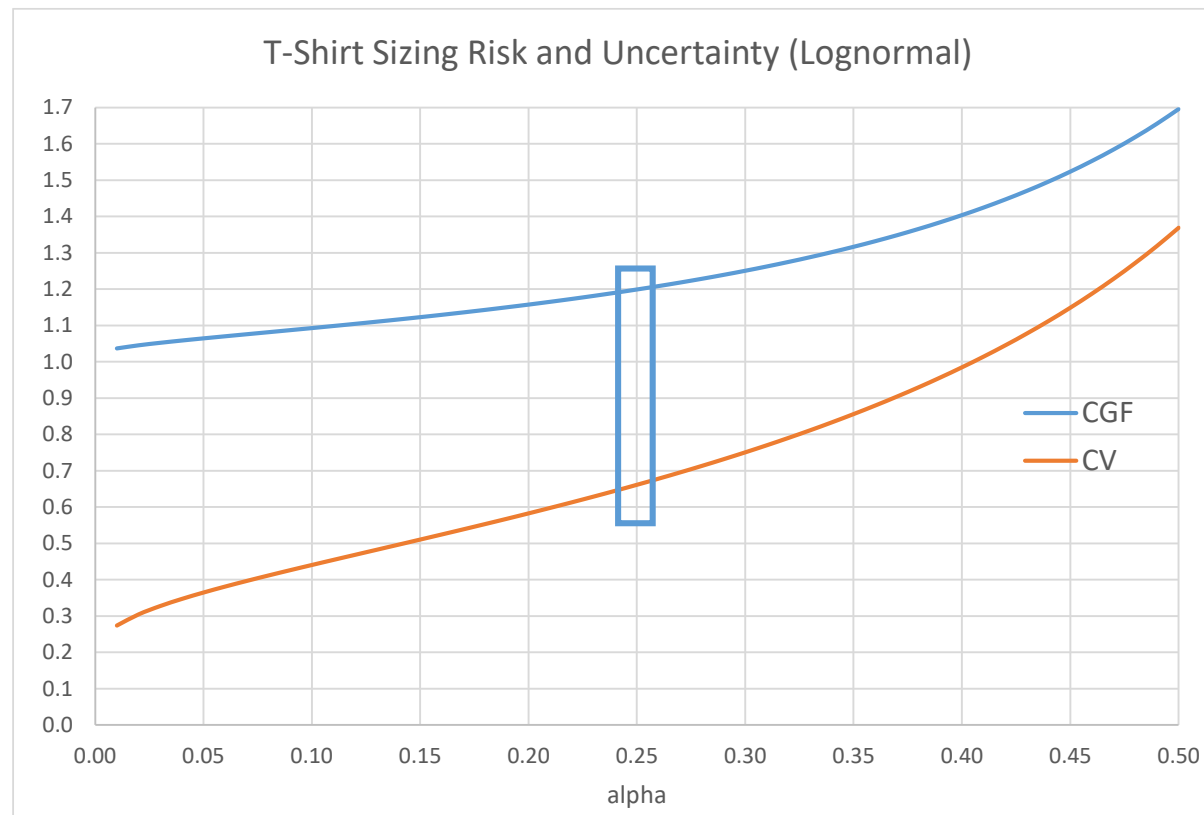
$$\sigma = \frac{\ln 2}{\Phi^{-1}(1 - \alpha/2)} = \frac{1}{\log_2 e^{\Phi^{-1}(1 - \alpha/2)}}$$

$$CV = \sqrt{e^{\sigma^2} - 1}$$



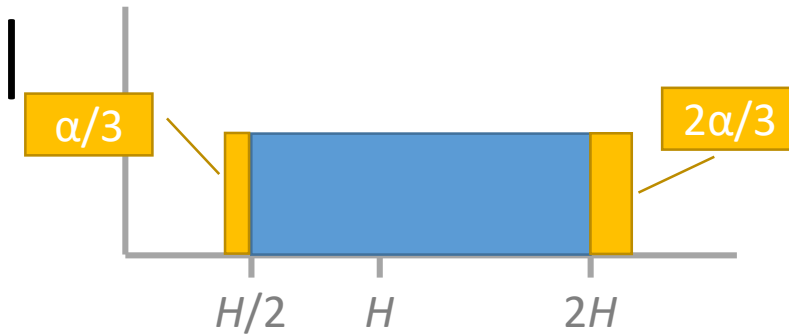
# T-Shirt Sizing Risk – Lognormal (Illustrated)

- Graph illustrates increase in CGF and CV as percent chance outside the “double-or-half” range increases
  - Beyond  $\alpha = 0.50$  (“coin flip”), values increase rapidly



# Uniform Expanded – Proportional

- Assume that the interval  $(H/2, 2H)$  encapsulates only  $(1 - \alpha)$  of the probability
  - That is, there is probability  $\alpha$  of being greater than  $2H$  or less than  $H/2$
  - This can be split proportionally or equally

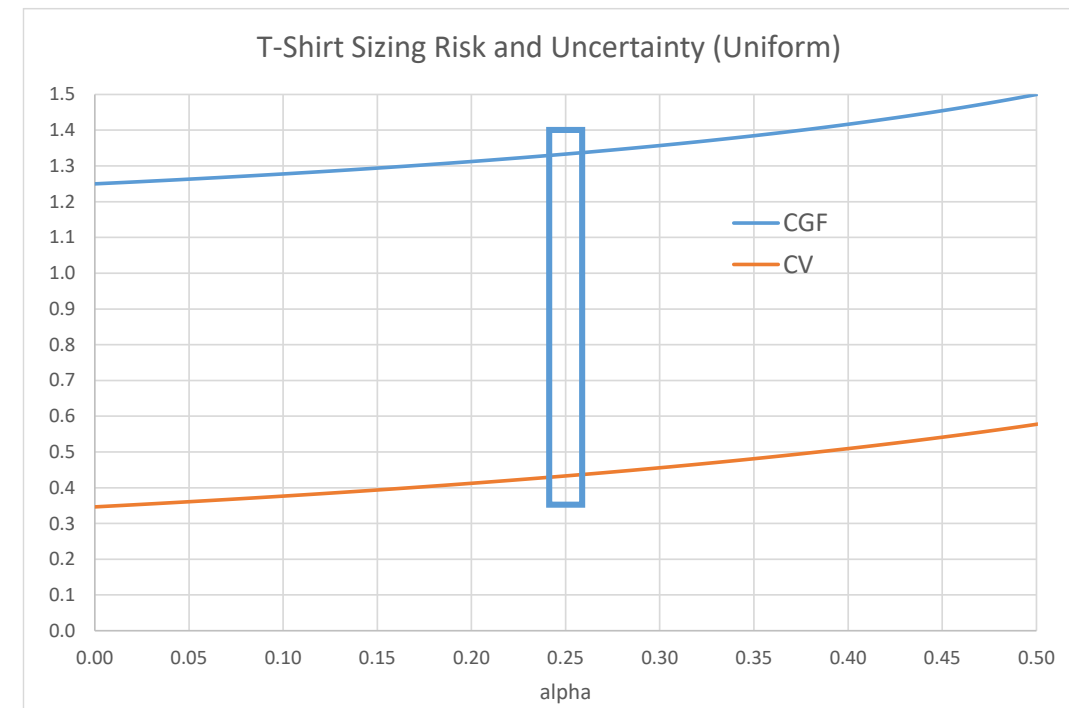


- Proportional puts  $\frac{2\alpha}{3}$  above and  $\frac{\alpha}{3}$  below

$$\mu = \left[ \frac{(1 - 2\alpha)H}{(1 - \alpha)} \frac{1}{2} + \frac{(2 - \alpha)H}{(1 - \alpha)} \right] / 2 = \frac{5 - 4\alpha}{4 - 4\alpha} H = \left( 1 + \frac{1}{4 - 4\alpha} \right) H$$

- Variance is range squared / 12:

$$\frac{(3H)^2}{12[2(1 - \alpha)]^2} = \left[ \frac{\sqrt{3}}{4 - 4\alpha} H \right]^2$$



# Uniform Expanded – Equal

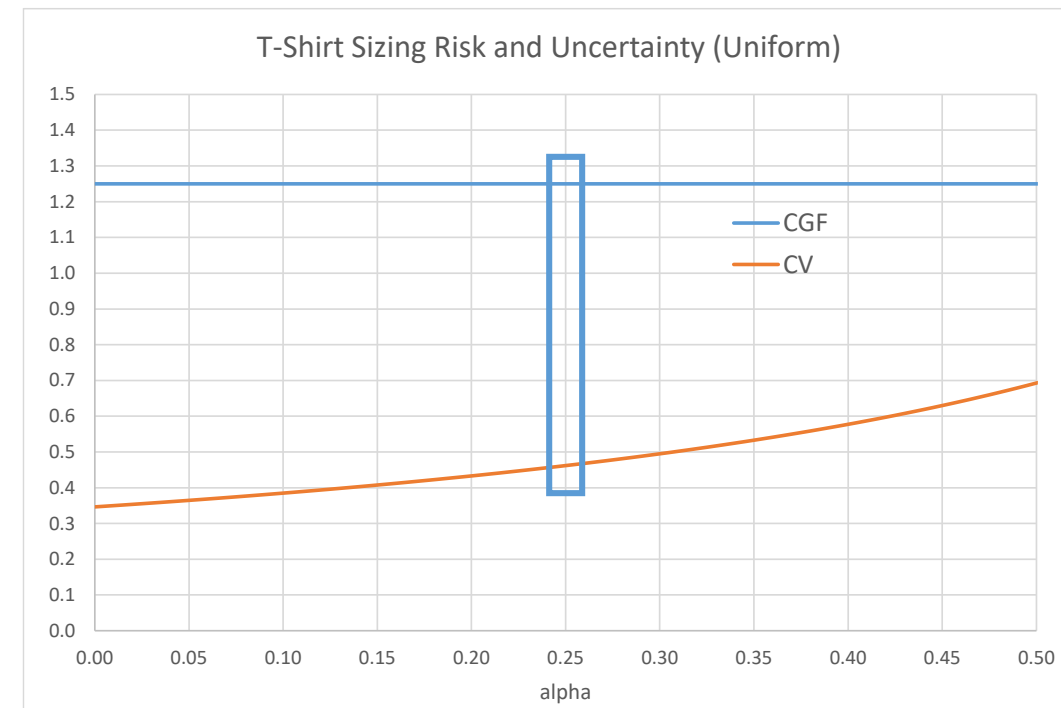
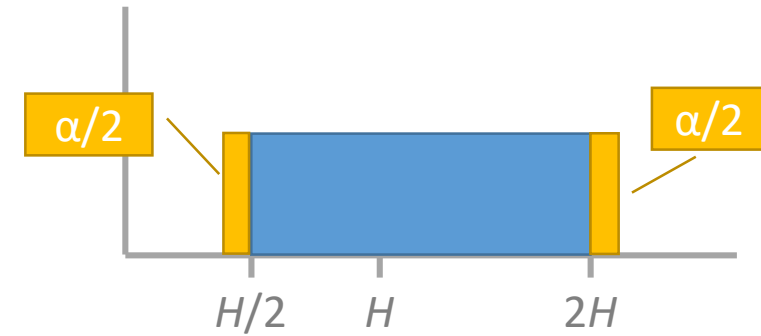
- Assume that the interval  $(H/2, 2H)$  encapsulates only  $(1-\alpha)$  of the probability
  - That is, there is probability  $\alpha$  of being greater than  $2H$  or less than  $H/2$
  - This can be split proportionally or equally

- Equal puts  $\frac{\alpha}{2}$  above and  $\frac{\alpha}{2}$  below

$$\mu = \left[ \frac{(2 - 5\alpha)}{(4 - 4\alpha)} H + \frac{(8 - 5\alpha)}{(4 - 4\alpha)} H \right] / 2 = \frac{5}{4} H = \left( 1 + \frac{1}{4} \right) H$$

- Variance is range squared / 12:

$$\frac{(6H)^2}{12[4(1 - \alpha)]^2} = \left[ \frac{\sqrt{3}}{4 - 4\alpha} H \right]^2$$



# Triangular Expanded – Proportional

- Assume that the interval  $(H/2, 2H)$  encapsulates only  $(1 - \alpha)$  of the probability

- That is, there is probability  $\alpha$  of being greater than  $2H$  or less than  $H/2$
- This can be split proportionally or equally

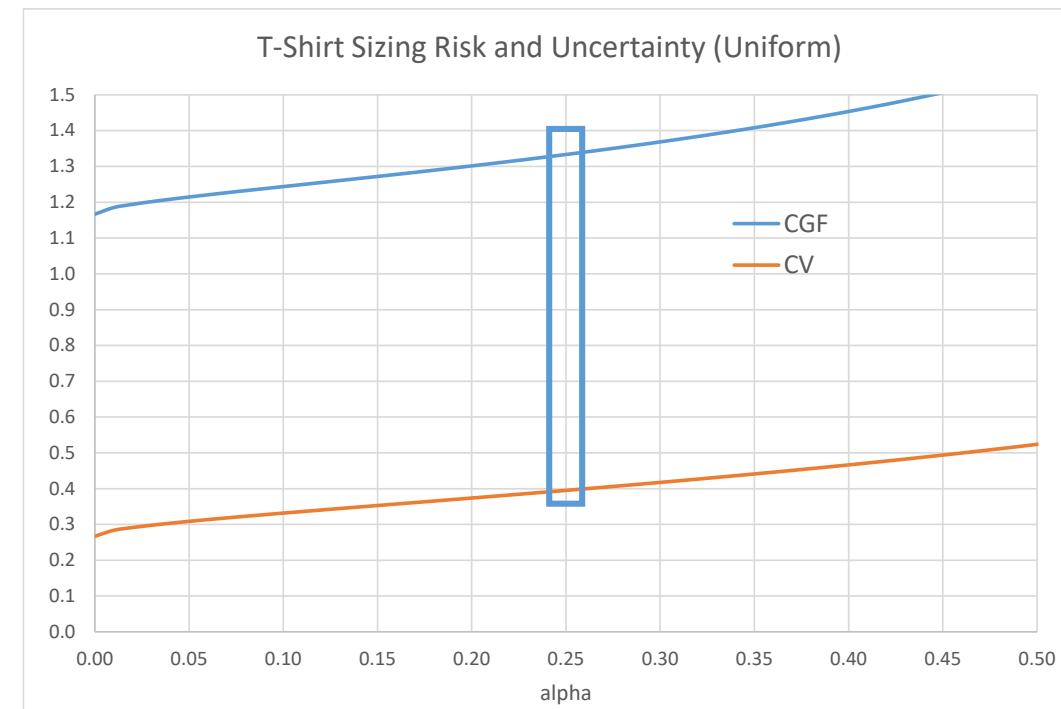
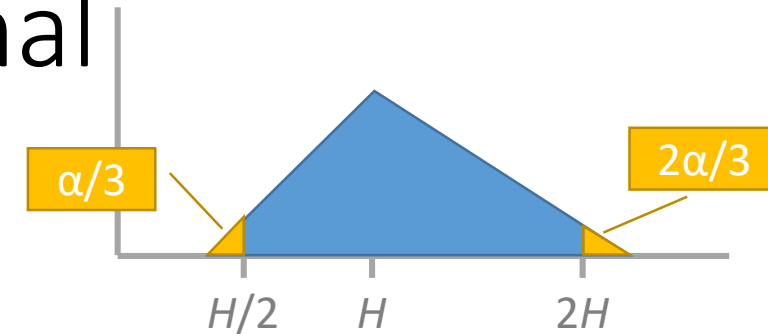
- Proportional puts  $\frac{2\alpha}{3}$  above and  $\frac{\alpha}{3}$  below

$$\mu = \left[ \left( 1 - \frac{\sqrt{\alpha}}{1 - \sqrt{\alpha}} \right) \frac{H}{2} + H + \left( 2 + \frac{\sqrt{\alpha}}{1 - \sqrt{\alpha}} \right) H \right] / 3 = \left( 1 + \frac{1}{6 - 6\sqrt{\alpha}} \right) H$$

- Variance:

$$\left[ \frac{\sqrt{\frac{7 - 4\sqrt{\alpha}}{2}}}{6 - 6\sqrt{\alpha}} H \right]^2$$

$$CV = \frac{\sqrt{\frac{7 - 4\sqrt{\alpha}}{2}}}{7 - 6\sqrt{\alpha}}$$



# Risk and Uncertainty by Confidence

- For confidence  $(1-\alpha)$ , we can express CGF and CV as a function of  $\alpha$ 
  - Generally, we would assume  $\alpha < 0.50$  (i.e., no worse than coin flip)

	Growth %	CV	Growth % ( $\alpha = 0.25$ )	CV ( $\alpha = 0.25$ )
Discrete (Generalized)	$\frac{\alpha}{4}$	$\frac{\sqrt{10\alpha - \alpha^2}}{4 + \alpha}$	6.2%	36.74%
Lognormal	$\sqrt{1 + CV^2} - 1$	$\sqrt{e^{\sigma^2}} - 1$	19.9%	66.16%
Uniform (Proportional)	$\frac{1}{4 - 4\alpha}$	$\frac{\sqrt{3}}{5 - 4\alpha}$	33.3%	43.30%
Uniform (Equal)	$\frac{1}{4}$	$\frac{\sqrt{3}}{5 - 5\alpha}$	25.0%	46.19%
Triangular (Proportional)	$\frac{1}{6 - 6\sqrt{\alpha}}$	$\frac{\sqrt{\frac{7 - 4\sqrt{\alpha}}{2}}}{7 - 6\sqrt{\alpha}}$	33.3%	39.53%

$$\sigma = \frac{1}{\log_2 e^{\Phi^{-1}(1-\alpha/2)}}$$

# Planning Poker and Fibonacci Numbers

- Alternate sizing method is Planning Poker
  - Commonly uses Fibonacci numbers for sizing via Story Points
  - In some alternative formulations, larger sizes are replaced with “rounder” numbers
  - Often visualized using fruits!
- Combines “additive” and “multiplicative” features:
  - Sum of any two consecutive sizes is equal to the next largest size
  - Ratio of consecutive sizes *approaches* a constant
- Fibonacci numbers are the sequence starting with 1 and 1, and whose subsequent entries are the sum of the two previous numbers
  - $2 = 1+1$ ,  $3 = 1+2$ ,  $5 = 2+3$ ,  $8 = 3+5$ ,  $13 = 5+8$ ,  $21 = 8+13$ ,  $34 = 13+21$ , etc.

# Fibonacci Numbers and the Golden Ratio

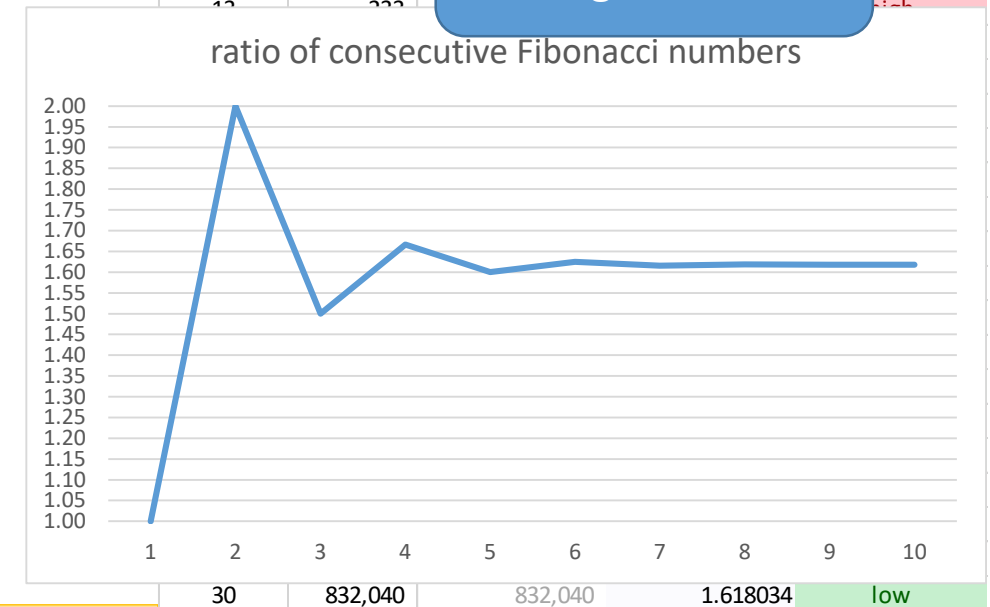
- Because the Fibonacci sequence is additive, the ratio between consecutive terms is not constant
- However, the ratio does quickly converge to a constant
  - It turns out that this is the Golden Ratio!

$$\phi = \frac{1 + \sqrt{5}}{2} = 1.618 \dots$$

$$F_n = \frac{1}{\sqrt{5}} [\phi^n - (1 - \phi)^n]$$

n	Fn	closed form	ratio	low/high
1	1	1		
2	1	1	1.000000	low
3	2	2	2.000000	high
4	3	3	1.500000	low
5	5	5	1.666667	high
6	8	8	1.600000	low
7	13	13	1.625000	high
8	21	21	1.615385	low
9	34			high
10	55			low
11	89			high
12	144			low

Factor = 1.618:1  
Range = 144:1

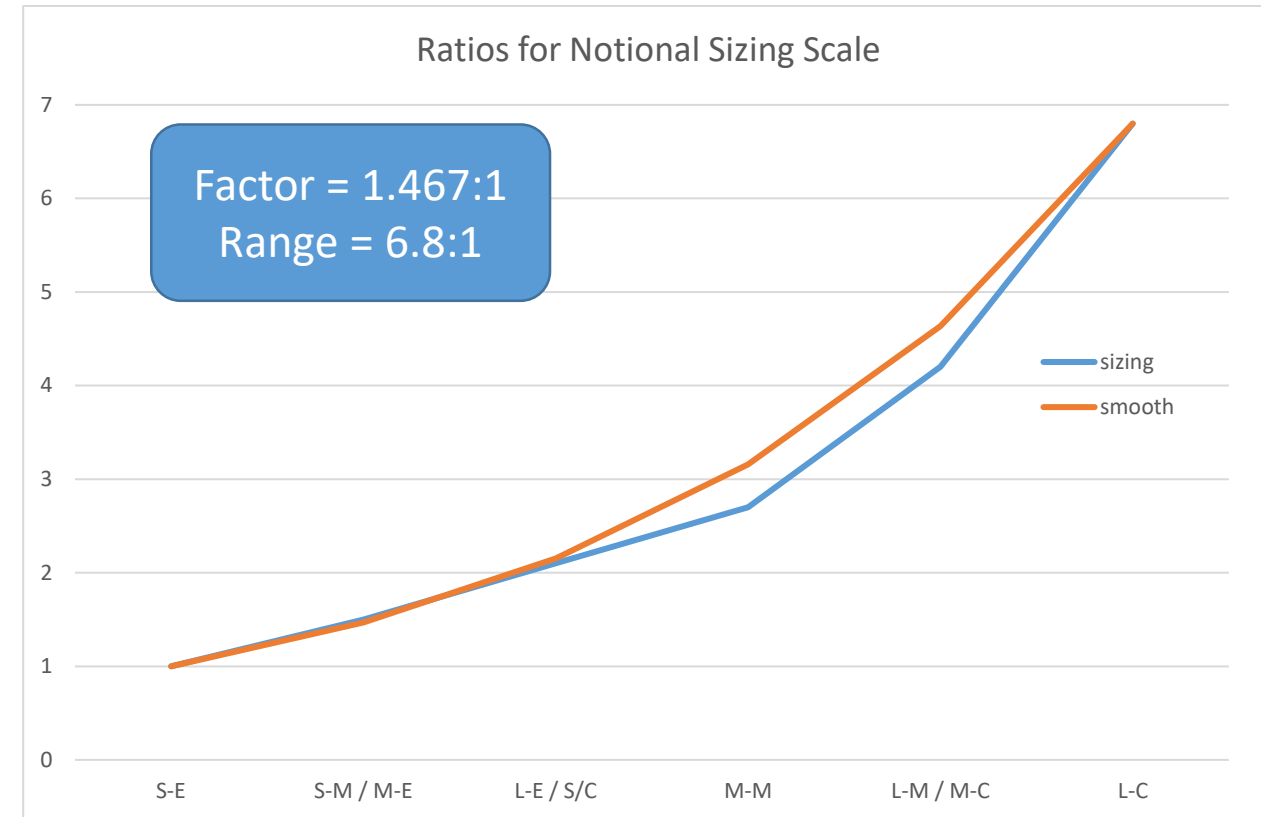


*The Golden Ratio: The Story of PHI, the World's Most Astonishing Number, Mario Livio, Crown, 2008.*



# Notional Sizing Model

- Incorporates Size and Complexity
  - Small, Medium, Large
  - Easy, Moderate, Complex
- Additional assumption of symmetry maps 3 x 3 model to 6-point scale
  - Total range 1 : 6.8
  - Average “notch” ratio 1.467



Sked (mo)	S	M	L
E	12	15	18
M	15	18	21
C	18	21	24

LOE (FTE)	S	M	L
E	2.5	3	3.5
M	3	4.5	6
C	3.5	6	8.5

effort (PM)	S	M	L
E	30	45	63
M	45	81	126
C	63	126	204

effort (relative)	S	M	L
E	37.0%	55.6%	77.8%
M	55.6%	100.0%	155.6%
C	77.8%	155.6%	251.9%

# Self-Similar Scales and the Ideal Ratio

- Self-similar scales are fractal in that misestimation will result in growth (or reduction) by the same ratio regardless of position on the scale
- Candidate ratios (R):
  - Two (2.0) – T-shirt Sizing
  - Phi (1.618...) – Planning Poker (Fibonacci numbers)
  - e (2.718...) – base of the exponential function that is its own derivative!
- It is proposed that these approximately bound the reasonable set of choices
- Related issue is “top-down” vs. “bottom-up”
  - Size more complex pieces of work as whole (initially) or force decomposition

# Empirical Testing of Scales

- Approach used in previous paper on use of SME's in Cost and Risk
  - Both knowable but unknown past events (e.g., box office gross of *Avengers: Endgame*) and unknown future events (e.g., box office gross of new release)
- Instead of asking for three-point estimates, ask for single best guess (closest value) from self-similar scale
  - Does gradation of scale affect accuracy of assessments?
- Expertise in subject area vs. expertise in uncertainty assessments

“Teaching Pigs to Sing: Improving Fidelity of Assessments from Subject Matter Experts (SMEs),”  
Peter Braxton and Richard Coleman, ICEAA Washington Chapter, June, 2012.

# Expert Judgment vs. Expert Opinion

- Expert Opinion = estimate is presented as a direct assessment by SME with no apparent basis
- Expert Judgment = SME uses or interprets data as the basis of the estimate, or at worst makes a direct assessment as to the scope on which the estimate is based (e.g., software sizing!)
- It is hypothesized that sizing and similar assessments can be improved by labeling each notch on the scale with an actual example reflecting that approximate size
  - Transcends Expert Opinion with a sort of a “stealth” Analogy
  - Heights of mountains, e.g., could be used in empirical assessment

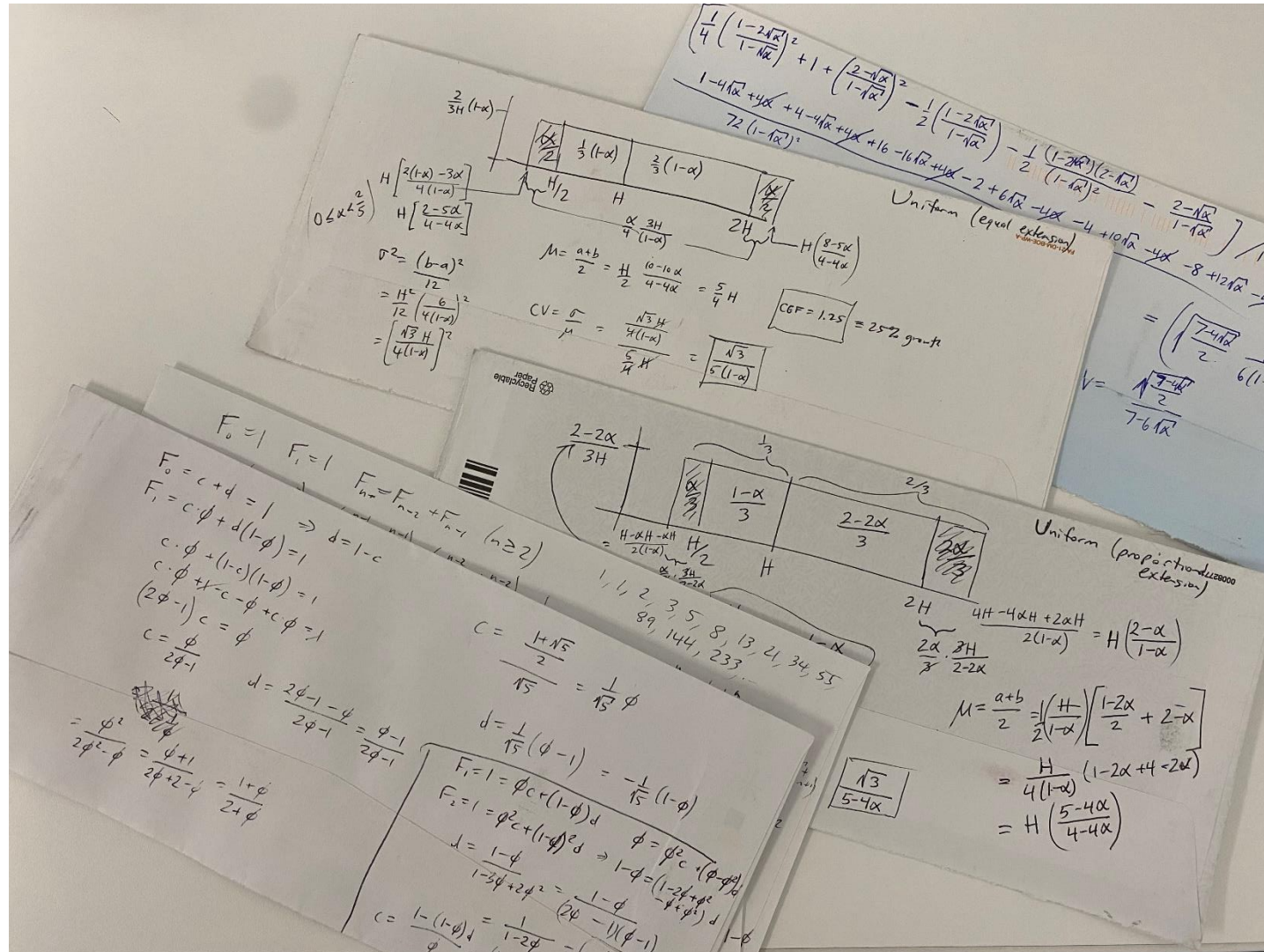
Cost Estimating Body of Knowledge (CEBoK®), Module 2 “Cost Estimating Techniques,” ICEAA, 2013.

# Conclusion

- More remains to be explored on empirical testing
- The bottom line is that significant risk and uncertainty are inherent in these self-similar sizing scales *even if we are off by no more than one size in either direction*

	Growth %	CV
Discrete ( $\alpha = 0.50$ )	12.5%	48.43%
Uniform	25.0%	34.64%
Triangular	16.7%	26.73%
Discrete ( $\alpha = 0.25$ )	6.3%	36.74%
Lognormal ( $\alpha = 0.25$ )	19.9%	66.16%
Uniform (Proportional)	33.3%	43.30%
Uniform (Equal)	25.0%	46.19%
Triangular (Proportional)	33.3%	39.53%

# Coda – The Proverbial Cocktail Napkin(s)



# Inherent Risk and Uncertainty of Self-Similar Sizing Scales in Agile Software Development

Back-Up

# Fibonacci Numbers Closed-Form Formula

- A closed-form formula can be derived, which will easily demonstrate the convergence property

- Suppose a relationship of the form

$$F_n = c \cdot a^n + d \cdot b^n$$

- Then the recursive formula will be satisfied if  $a$  and  $b$  are roots of the quadratic

$$F_n + F_{n+1} = c \cdot a^n + d \cdot b^n + c \cdot a^{n+1} + d \cdot b^{n+1}$$

$$= c(a^n + a^{n+1}) + d(b^n + b^{n+1}) = c \cdot a^{n+2} + d \cdot b^{n+2} = F_{n+2}$$

$$x^2 = x + 1 \rightarrow x^2 - x - 1 = 0 \rightarrow a = \frac{1 + \sqrt{5}}{2} = \phi, b = \frac{1 - \sqrt{5}}{2} = 1 - \phi$$

- Now we solve for the coefficients  $c$  and  $d$

$$F_1 = 1 = \phi c + (1 - \phi)d, F_2 = 1 = \phi^2 c + (1 - \phi)^2 d$$

$$c = \frac{1}{2\phi - 1} = \frac{1}{\sqrt{5}}, d = \frac{1}{1 - 2\phi} = -\frac{1}{\sqrt{5}} \rightarrow F_n = \frac{1}{\sqrt{5}} [\phi^n - (1 - \phi)^n]$$

- Since the second term vanishes as  $n$  increases without bound, the ratio of consecutive terms approaches  $a$